Accepted Manuscript

Estimation of a room ventilation air change rate using a stochastic grey-box modelling approach

Marcel Macarulla, Miquel Casals, Núria Forcada, Marta Gangolells, Alberto Giretti

PII: S0263-2241(18)30303-8
DOI: https://doi.org/10.1016/j.measurement.2018.04.029
Reference: MEASUR 5430

To appear in: *Measurement*

Received Date: 5 October 2017
Revised Date: 3 April 2018
Accepted Date: 9 April 2018

Estimation of a room ventilation air change rate using a stochastic grey-box modelling approach

Marcel Macarulla*[a], Miquel Casals[a], Núria Forcada[a], Marta Gangolells[a], Alberto Giretti[b]

Authors' affiliation

[a]Universitat Politècnica de Catalunya, Department of Project and Construction Engineering, Group of Construction Research and Innovation (GRIC), C/Colom, 11, Ed. TR5, 08222 Terrassa, Barcelona, Spain; [b]Università Politecnica delle Marche, Department of Civil and Building Engineering and Architecture, Via delle Brecce Bianche, 60100 Ancona, Italy

Corresponding author

Marcel Macarulla, PhD, Universitat Politècnica de Catalunya, Department of Project and Construction Engineering, Group of Construction Research and Innovation (GRIC), C/Colom, 11, Ed. TR5, 08222 Terrassa, Barcelona, Spain. Tel: (+34) 937398395, marcel.macarulla@upc.edu

Highlights

The grey-box modelling approach is used to estimate the room ventilation air change rate

Four days are enough to obtain suitable parametrization with grey-box modelling

The CO_2 sensor location is not important when the grey-box modelling approach is used

Abstract

The measurement of ventilation air change rate is a difficult, expensive task in buildings. Usually, the tracer-gas mass balance equation is used to determine ventilation air change rates. This method uses an ordinary differential equation. Consequently, it cannot deal with disturbances that enter the system, such as the influence of unrecognized and unmodelled inputs or the measurement noise. The use of the stochastic grey-box modelling approach, which is less common in the ventilation field, can help to deal with disturbances that can affect the system. The objective of this paper is to assess the potential of using the stochastic grey-box modelling approach to estimate the ventilation air change rate. The modelling is based on the stochastic differential equation of tracer-gas mass balance. The results show that this approach produces robust estimations to determine the ventilation air change rate of a room.

Keywords: indoor air quality, ventilation, air change rate estimation, stochastic methods, low-order model

1 Introduction

During a building's operation, one of the aims is to ensure good indoor air quality: i.e., a healthy and comfortable indoor environment [1]. This goal should be attained using low-energy solutions [2]. In this context, knowledge of the ventilation air change rate is crucial to efficiently monitor buildings' indoor air quality [3].

The measurement of ventilation air change rate is a difficult, expensive task [4]. Currently, CO_2 concentration is widely used as a tracer gas to estimate air change rates in the building research field [3,5,6]. This method uses the ordinary differential equation (ODE) of the tracer-gas mass balance to determine the air change rates. The approach is deterministic, because ODE solutions are deterministic functions of time. Consequently, it is assumed that future concentrations and effects can be predicted exactly [7].

Measurement practice generally relies on just one sensor in a single room [3,8] and it is assumed in most studies that the air in the space is perfectly mixed [9]. However, as much research has demonstrated, this assumption is not true. In a specific space, there is a spatial distribution of carbon dioxide, influenced by the occupants' location and movement, the heat sources, and the room air distribution [5,10]. Consequently, extreme care should be taken in the positioning of the CO_2 sensor [9]. The measurement of CO_2 concentration at a single location or height may not act as a true representation of the CO_2 concentration in a space [10]. Therefore, the estimation of air change rate in a specific room could vary depending on the sensor location when deterministic approaches are used.

Deterministic approaches cannot deal with disturbances that enter the system, such as the influence of unrecognized, unmodelled inputs or measurement noise [11,12]. Hence, the deterministic approach cannot manage uncertainties that affect the system.

The use of the stochastic approach can help to deal with disturbances that may affect the system. Stochastic differential equation (SDE)-based models, also known as grey-box models, combine the physical knowledge of a system and the information embedded in the monitoring data [13]. The stochastic differential equations can be written as:

$$dX_t = f(X_t, U_t, \theta, t)dt + G(\theta, t)dW_t \qquad \text{(Equation 1)}$$

$$Y_{t_k} = h(X_t, U_t, \theta, t) + e_t \qquad \text{(Equation 2)}$$

Equations describing the physical aspects of the system X_t are formulated in continuous time and composed of a drift term, $f(X_t, U_t, \theta, t)dt$ and a diffusion term, $G(\theta, t)dW_t$. The diffusion term represents the approximations and the noise introduced into the system due to unknown or unmodelled disturbances. The diffusion term is composed of a function describing how the disturbance enters the system, $G(\theta, t)$, and a standard Wiener process W_t. The discrete time observations, Y_{t_k}, include the measurement error e_t that is assumed to be Gaussian distributed. The measurement error represents the noise due to the light inaccuracy of the sensors used. U_t represents the inputs and θ the parameters of the system. In the literature, the maximum likelihood method is generally used to determine θ[11,12].

If the residual error is separated into diffusion and measurement noise, the model can be validated, because a more accurate description of the prediction error is obtained. If the model is detailed enough to describe the dynamics of the system, the residuals will be uncorrelated [7]. Furthermore, the inclusion of the diffusion term helps to determine how to improve an insufficient model [14].

The objective of this paper is to show how grey-box modelling can be used to estimate the ventilation air change rate in a room. In addition, the aim is to discuss the number of CO_2 sensors and the distribution required to make this estimation. The paper is divided into four sections. Section 2 describes the methodology used in this research. In Section 3, the results are presented and discussed. Finally, Section 4 presents the conclusions.

2 Methodology

The methodology used in this paper to determine ventilation air change rate in a chamber is divided into 4 steps: experimental design and data collection, modelling process, validation process, and comparison of ventilation air change rates obtained with the models (Fig. 1).

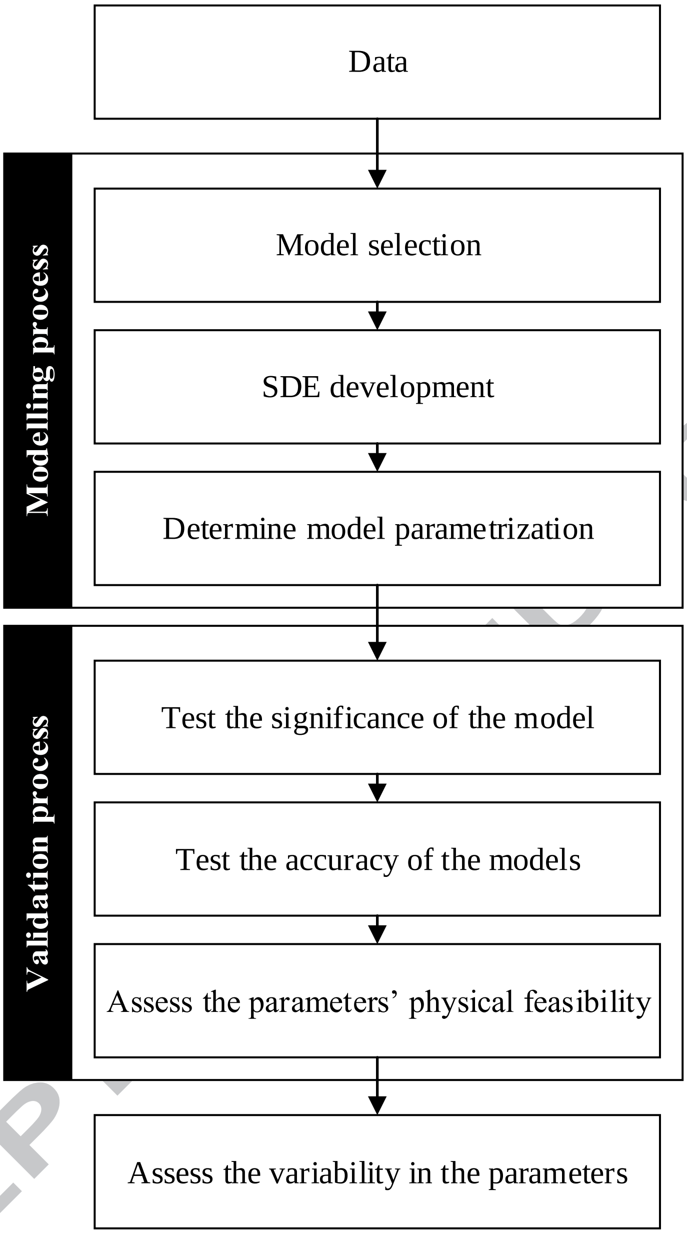

Fig. 1. Methodology

2.1 Experimental design and data collection

The room under study is an office in an academic building in Terrassa (Barcelona). It contains six work stations that are used by PhD and master's degree students. In other words, the main users of this room are university students. Although the maximum occupancy is six people, during the experiment the maximum occupation reached was three people.

The room is 3.05 m high, with a surface area of 31.16 m^2 and a volume of 95 m^3. It is accessed through a door located in a corridor. It has one aluminium window, which is well-conserved, 1.40 m long, 1.70 m high, and 1.25 m above floor level. The room is ventilated by a ventilation grille in the access door. The ventilation grille is 0.40 m long, 0.20 m high, and 0.10 m above floor level. The room has neither forced ventilation nor heating or cooling sources. The window was closed throughout the test period.

The CO_2 concentration in the room was measured using two Advanticsys IAQM-THCO2 sensors with a range from 0 to 3,000 ppm, a resolution of 1 ppm, and an accuracy of ±2%, full scale. The sensors were calibrated and configured by the manufacturer to record an instantaneous value every 15 minutes.

The CO_2 concentration in an occupied room is not homogeneous, due to people breathing, infiltrations through the windows, and the influence of heating and cooling sources. According to Bulińska [3], CO_2 sensors should not be positioned in the areas around people, windows or radiators. The representative location is the centre of the room. In this research, four CO_2 sensors were used. One sensor was placed in the centre of the room under the ceiling (SC), 3.05 m above the floor. The second was located slightly away from the centre of the room at 1.20 m from the ground (SD) (see Fig. 2). Another CO_2 sensor was put in the corridor to measure the CO_2 concentration of the ventilation flow (C_{ven}). Finally, another sensor was located outside the building to measure the external CO_2 concentration (C_{ext}). Fig. 2 presents a plan view with all physical measurement conditions.

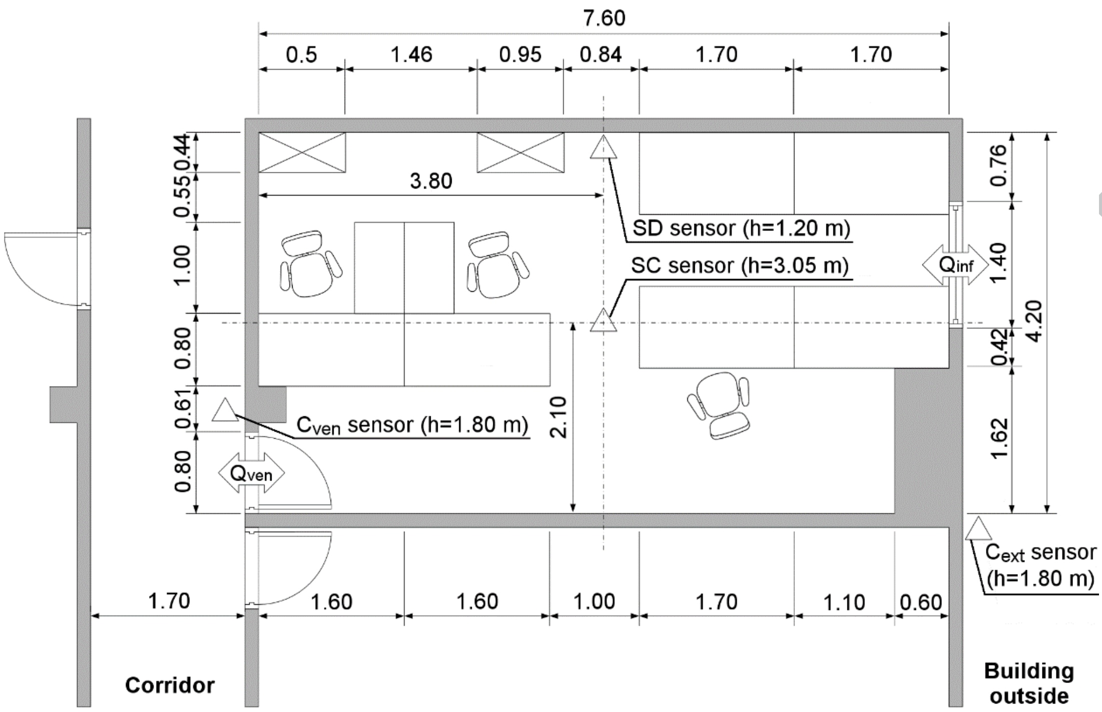

Figure 2: Plan view of the test room (all measurements are in metres)

Table 1. Measurement conditions

Location	Parameter	Units	Value
Corridor air parameters	Mean temperature	ºC	28.1
	Mean relative humidity	%	49
Room air parameters	Mean temperature	ºC	27.7
	Mena relative humidity	%	51
Outdoor air parameters	Mean temperature	ºC	26.0
	Mean relative humidity	%	58
	Mean atmospheric pressure	hPa	1019.6
	Mean wind speed	m/s	1.9

The occupancy of the room was obtained from an occupancy sheet. The occupancy sheet was used to calculate the mean occupancy every quarter of an hour. The testing period was carried out over four days in July 2016. Fig. 3 presents the data set used in this research, and Table 1 presents the indoor and outdoor air conditions during the experiment.

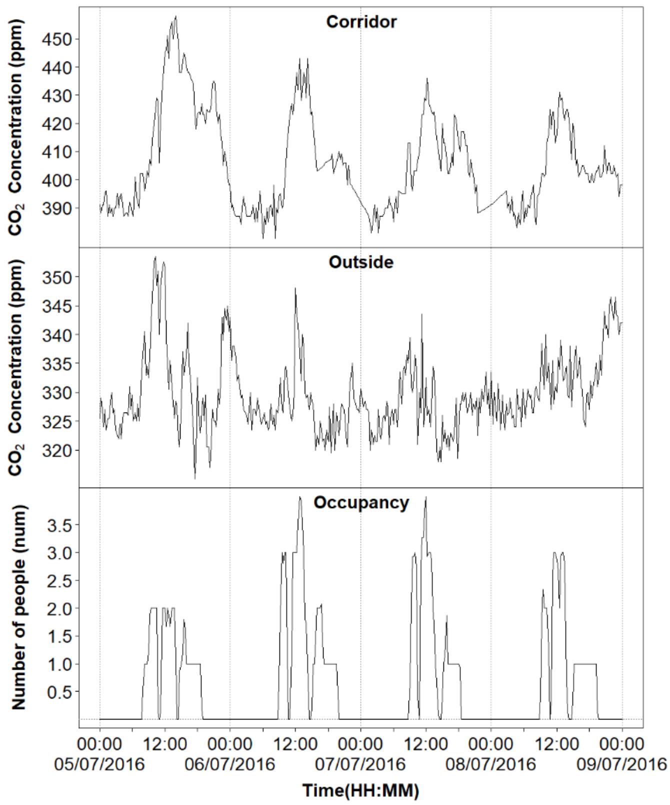

Fig. 3. Data set used for this research. From the top, the first plot shows the CO_2 concentration observed in the corridor, the second shows the CO_2 concentration observed outside the building, and the last plot presents the occupancy of the room.

2.2 Modelling process

Grey-box models combine physical knowledge of a system with information embedded in the monitored data [13]. The physical knowledge of the system is described as a set of first-order stochastic differential equations, composed of a drift term and a diffusion term of the process. The drift term is a function that describes the well-known physics of the system. In this study, the tracer-gas mass balance [15–19] was used as a drift term

(Eq. 3), representing the change in CO_2 concentration (C_{int}) at a point in time in a room with volume V_r and two ventilation flows Q_{ven} and Q_{inf}:

$$\frac{dC_{int}}{dt}V_r = (C_{ven} - C_{int}) \cdot Q_{ven} + (C_{ext} - C_{int}) \cdot Q_{inf} + G_{CO_2} \quad \text{(Eq. 3)}$$

where C_{ven} is the CO_2 concentration of the ventilation rate, Q_{ven} is the ventilation flow from the corridor, Q_{inf} are the window infiltrations, C_{ext} is the CO_2 concentration from the exterior of the building, and G_{CO_2} is the CO_2 generated by the occupants.

Equation 3 assumes that CO_2 is chemically stable and inert, and there is no absorption process that can reduce the CO_2 concentration. Hence, walls, ceiling and furniture do not absorb CO_2. Finally, the above equation assumes a perfectly mixed condition, and constant ventilation air flows [19]. In this research, the two ventilation air flows that were considered were constant natural air flows. G_{CO_2} was calculated using the following equation:

$$G_{CO_2} = K_{occ} \cdot P \quad \text{(Eq. 4)}$$

where K_{occ} is the CO_2 exhaled per occupant, and P is the occupancy of the room. Equation 4 assumes that the CO_2 exhaled per occupant is constant over time and the same for each person. This is a reasonable hypothesis, because the occupants carried out normal office activities: sitting and reading or writing. The increase in CO_2 exhalation due to people walking around when they entered or left the room was assumed to be negligible.

To complete the stochastic differential equation, the diffusion terms are added:

$$dC_{int} = \frac{Q_{ven}}{V_r}(C_{ven} - C_{int}) \cdot dt + \frac{Q_{inf}}{V_r}(C_{ext} - C_{int}) \cdot dt + \frac{K_{occ} \cdot P}{V_r} \cdot dt + \sigma \cdot dw \quad \text{(Eq. 5)}$$

where dw is a Wiener process, and σ is the incremental variance in the Wiener process. Eq. 5 is a stochastic differential equation that represents the behaviour of the CO_2 concentration over time. This equation can be represented using a RC-network, as shown in Fig. 4.

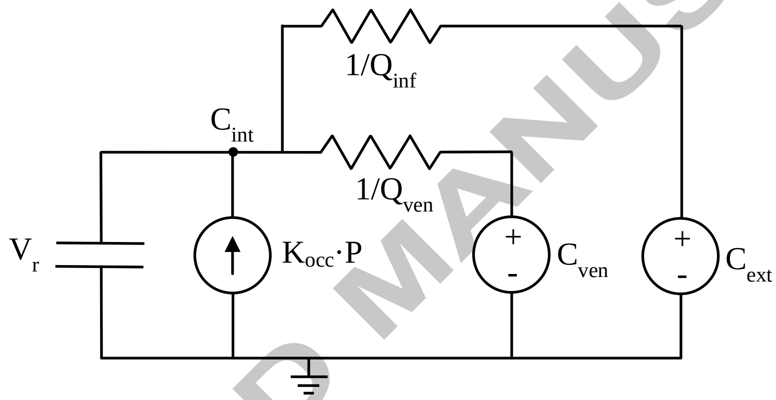

Fig. 4. RC-network of the model described by Eq. 5

Finally, the monitored output of the system was defined in this research using the following discrete time equation:

$$Y_{t_k} = C_{int,\ t_k} + e_k \quad \text{(Eq. 7)}$$

where Y_{t_k} is the measured interior CO_2 concentration of the room at time t_k, and e_k is a white noise process that describes the noise of the measurements.

In this study, three models were used to represent the interior CO_2 concentration of the room (see Table 2), based on Eq. 3. Model 1 (M1) considered that the ventilation flow due to window infiltrations was negligible, and the CO_2 concentration of the ventilation air flow constant had a value of 404 ppm [20]. The aim was to test whether it was possible to estimate the ventilation air change rate of a room using only sensors located inside the room. This model used the occupancy of the room as input.

Model 2 (M2) also considered that the ventilation flow due to infiltrations was negligible. However, in this case, a CO_2 sensor located in the ventilation flow was used. The aim was to test whether the room ventilation air change rate could be estimated with one sensor in the ventilation flow and sensors inside the room. The inputs used in this model were occupancy and CO_2 concentration of the ventilation air flow.

Finally, Model 3 (M3) considered the full model proposed in Equation 3. The aim was to test whether better estimations could be made using a greater number of sensors. This model required one sensor in the ventilation air flow, one sensor outside the building, and sensors inside the room. Occupancy and the CO_2 concentration of the ventilation air flow were used as inputs.

The room CO_2 concentration was measured by two sensors, resulting in three readings used as observations in the aforementioned models: Reading 1 (R1) corresponds to the reading produced by the CO_2 sensor located on the ceiling; Reading 2 (R2) corresponds to the reading produced by the CO_2 sensor located on the desk; and Reading 3 (R3) is the mean value of the CO_2 sensor located on the desk and the CO_2 sensor located on the ceiling. With the above observations and inputs for each model, we could estimate the unknown parameters of each model. All models used the same initial values and bounds for each parameter for the estimations. The initial values and bounds used for each parameter are presented in Table 3.

Table 2. Characteristics of the assessed models

Model-Reading	State space representation	Observation	Input	Estimated parameters
M1-R1	$dC_{R1} = \frac{Q_{\text{ven}}}{95}(404 - C_{R1}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{\text{R1},\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R1}	P	C_{R1} Q_{ven}, K_{occ}, σ, σ_{e_k}
M1-R2	$dC_{R2} = \frac{Q_{\text{ven}}}{95}(404 - C_{R2}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{R2,\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R2}	P	C_{R2}, Q_{ven}, K_{occ}, σ, σ_{e_k}
M1-R3	$dC_{\text{R3}} = \frac{Q_{\text{ven}}}{95}(404 - C_{\text{R3}}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{\text{R3},\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R3}	P	C_{R3}, Q_{ven}, K_{occ}, σ, σ_{e_k}
M2-R1	$dC_{R1} = \frac{Q_{\text{ven}}}{95}(C_{\text{ven}} - C_{R1}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{R1,\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R1}	P, C_{ven}	C_{R1} Q_{ven}, K_{occ}, σ, σ_{e_k}
M2-R2	$dC_{R2} = \frac{Q_{\text{ven}}}{95}(C_{\text{ven}} - C_{R2}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{R2,\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R2}	P, C_{ven}	C_{R2}, Q_{ven}, K_{occ}, σ, σ_{e_k}
M2-R3	$dC_{\text{R3}} = \frac{Q_{\text{ven}}}{95}(C_{\text{ven}} - C_{\text{R3}}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{\text{R3},\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R3}	P, C_{ven}	C_{R3}, Q_{ven}, K_{occ}, σ, σ_{e_k}
M3-R1	$dC_{\text{R1}} = \frac{Q_{\text{ven}}}{95}(C_{\text{ven}} - C_{R1}) \cdot dt + \frac{Q_{\text{inf}}}{95}(C_{ext} - C_{R1}) \cdot dt + \frac{K_{\text{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{\text{R1},\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R1}	P, C_{ven}, C_{ext}	C_{R1}, Q_{ven}, Q_{inf}, K_{occ}, σ, σ_{e_k}

M3-R2	$dC_{R2} = \frac{Q_{\mathrm{ven}}}{95}(C_{\mathrm{ven}} - C_{R2}) \cdot dt + \frac{Q_{\mathrm{inf}}}{95}(C_{ext} - C_{\mathrm{R2}}) \cdot dt + \frac{K_{\mathrm{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{R2,\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R2}	P, C_{ven}, C_{ext}	C_{R2}, Q_{ven}, Q_{inf}, K_{occ}, σ, σ_{e_k}
M3-R3	$dC_{R3} = \frac{Q_{\mathrm{ven}}}{95}(C_{\mathrm{ven}} - C_{R3}) \cdot dt + \frac{Q_{\mathrm{inf}}}{95}(C_{ext} - C_{R3}) \cdot dt + \frac{K_{\mathrm{occ}} \cdot P}{95} \cdot dt + \sigma \cdot dw$ $Y_{t_k} = C_{\mathrm{R3},\, t_k} + \sigma_{e_k} \cdot e_k$	C_{R3}	P, C_{ven}, C_{ext}	C_{R3}, Q_{ven}, Q_{inf}, K_{occ}, σ, σ_{e_k}

Table 3. Initial values

Parameter	Intial value	Lower bound	Upper bound
C	404	0	3000
Q_{ven}	50	0	500
Q_{inf}	50	0	500
K_{occ}	42000	1000	90000
σ	Exp(1)	Exp(-20)	Exp(20)
σ_{e_k}	Exp(1)	Exp(-50)	Exp(20)

2.3 Model validation

To test the significance of the calculated ventilation air change rate, the first step was to validate all the models used to obtain the parametrizations. Grey-box models can be validated using a set of statistical tests [11,13]. The statistical tests used in this research to validate the models were based on previous literature [11,13,21].

First, the significance of all model parameters was assessed. For this purpose, the t-test scores and associated p-values were calculated for each parameter. The p-value of all parameters should be less than 0.05, otherwise the parameter is considered insignificant. Additionally, the correlation matrix of the estimated parameters was used to test for over-parameterization. The off-diagonal values of the correlation matrix should be far from 1 or -1. If the off-diagonal values are near to 1 or -1, the model is over-parameterized [22].

To test whether the solution was the true optimum, the derivative of the objective function with respect to the particular initial state or parameter was assessed. This value should be close to zero if the solution is the true optimum [23].

The derivative of the penalty function with respect to the particular initial state or parameters was used to test whether the initial state or parameter was close to one of its limits. This value should not be significant when it is compared to the derivative of the objective function with respect to the particular initial state or parameter. If the value is significant compared to the derivative of the objective function, the particular initial state or parameter may be close to one of its limits. Then, new limits should be set and the estimation should be repeated [22].

Residuals of the pure simulation were used to evaluate the model accuracy. In addition, one-step-ahead residual analysis can be used to assess the assumption of white noise. For this purpose, the autocorrelation function and the cumulated periodogram were used [11,13]. The accuracy of the models was assessed using the root mean square error. All the above statistical tests were calculated using the entire dataset and were provided by the CTSM-R package [22].

Finally, an analysis of physical feasibility of the estimated parameters was carried out. The estimated parameters were compared with similar studies to assess whether they were consistent with reality.

2.4 Variability of the estimations

This study used three models with three different readings to determine the ventilation air change rate. In addition, the human emission rate of CO_2 was estimated. First, the 95% confidence interval for each estimation was calculated. The aim was to determine statistically whether the estimations of each parameter were equal to each other. Finally, the coefficient of variation for each parameter was calculated to determine the variability between estimations.

3 Results and discussion

This section presents the validation results for the three proposed models and for each position. The validation process is described in Section 3.2 and is divided into three steps. The significance of the model was assessed, its accuracy was tested, and the physical feasibility of the estimated parameters was evaluated. Finally, the variability of the estimated parameters was discussed to define the best configuration to determine the ventilation rate of a building space. Table 4 summarizes the results of the validation process.

Table 4. Summary of the results for each model

		M1-R1	M1-R2	M1-R3	M2-R1	M2-R2	M2-R3	M3-R1	M3-R2	M3-R3
Model significance	Model overparametrized?	No	No	No	No	No	No	Yes	Yes	Yes
	True optimum?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	White noise?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model accuracy	RMSE (ppm)	40	45	39	35	44	36	35	44	36
Physical feasibility of the model parameters	Q_{ven} estimation (m^3/h)	34.29 ($0.36\ h^{-1}$)	28.77 ($0.30\ h^{-1}$)	31.26 ($0.33\ h^{-1}$)	35.31 ($0.37\ h^{-1}$)	29.32 ($0.31\ h^{-1}$)	32.02 ($0.34\ h^{-1}$)	35.31 ($0.37\ h^{-1}$)	29.32 ($0.31\ h^{-1}$)	32.02 ($0.34\ h^{-1}$)
	Q_{ven} standard error (m^3/h)	0.9 ($0.01\ h^{-1}$)	1.13 ($0.01\ h^{-1}$)	0.83 ($0.01\ h^{-1}$)	0.89 ($0.01\ h^{-1}$)	1.16 ($0.01\ h^{-1}$)	0.86 ($0.01\ h^{-1}$)	0.82 ($0.01\ h^{-1}$)	1.13 ($0.01\ h^{-1}$)	0.83 ($0.01\ h^{-1}$)
	Q_{ven} estimation variability (%)	2.62	3.94	2.64	2.52	3.94	2.69	2.33	3.86	2.58
	Q_{inf} estimation (m^3/h)	-	-	-	-	-	-	0.00 ($0.00\ h^{-1}$)	0.00 ($0.00\ h^{-1}$)	0.00 ($0.00\ h^{-1}$)
	Q_{inf} standard error (m^3/h)	-	-	-	-	-	-	0.00 ($0.00\ h^{-1}$)	0.00 ($0.00\ h^{-1}$)	0.00 ($0.00\ h^{-1}$)
	Q_{inf} estimation variability (%)	-	-	-	-	-	-	0.00	0.00	0.00
	K_{occ} estimation (L CO^2/h)	14.18	13.62	13.83	14.18	13.56	13.80	14.18	13.56	13.80
	K_{occ} standard error (L CO^2/h)	0.28	0.42	0.30	0.26	0.42	0.27	0.25	0.40	0.27
	K_{occ} estimation variability (%)	1.97	3.07	2.14	1.84	3.06	1.96	1.75	2.94	1.93
Model computation time	Time (s)	2.61	2.35	2.44	2.51	2.43	2.33	2.97	2.61	2.48

All the statistical tests calculated for M1 and M2 for all the readings reported similar results. The p-values of the t-tests were below 0.05 for all estimated parameters. Values on the off-diagonal of the above models were not close to 1 or -1. Therefore, these models were not over-parametrized. M3-R1, M3-R2 and M3-R3 p-values of the t-tests were below 0.05 for all parameters, except for the estimate of Q_{inf} (M3-R1: Q_{inf} p-value=0.996; M3-R2: Q_{inf} p-value=0.996; M3-R3: Q_{inf} p-value=0.995). This means there is no evidence that Q_{inf} was different from 0.

The derivative of the objective function with respect to each parameter for all models and readings was close to 0. The derivative of the penalty function with respect to the particular initial state or parameter was not significant compared with the derivative of the objective function in all models and readings. Therefore, the solutions found could be the true optimum and the solutions were not close to the limits.

An analysis of the pure simulation residuals revealed that those of M1 were slightly higher (Fig. 5) than residuals of M2 and M3 (Fig. 6 and Fig. 7). The shape of residuals was similar in all models, however the amplitude of the residuals in M1 was slightly higher than in M2 and M3. This could be clearly observed on the first day of the testing period, and was due to the fact that M1 did not consider the CO_2 concentration of the ventilation air flow. This effect was magnified when the variability of the CO_2 concentration of the ventilation air flow increased. Other studies reported the same effect (see [19]). According to these results, the use of one sensor in the ventilation air flow improves the performance of the pure simulation residuals.

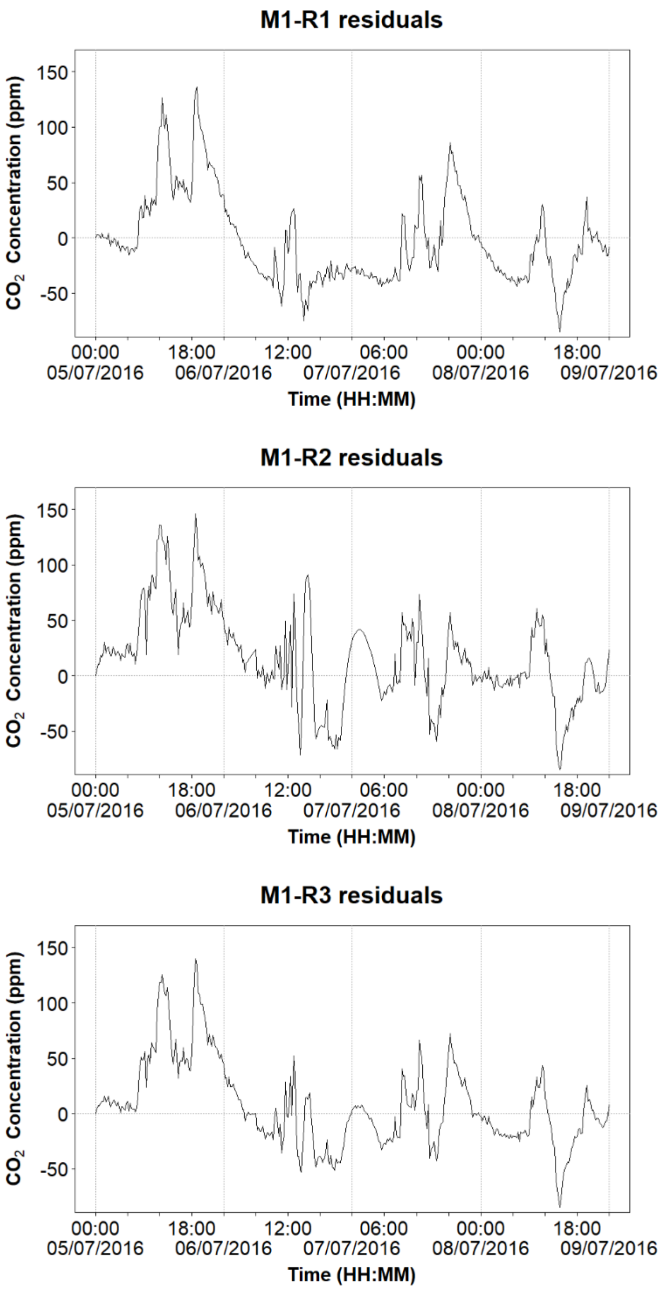

Fig. 5. Pure simulation residual plots for M1-R1, M1-R2 and M1-R3.

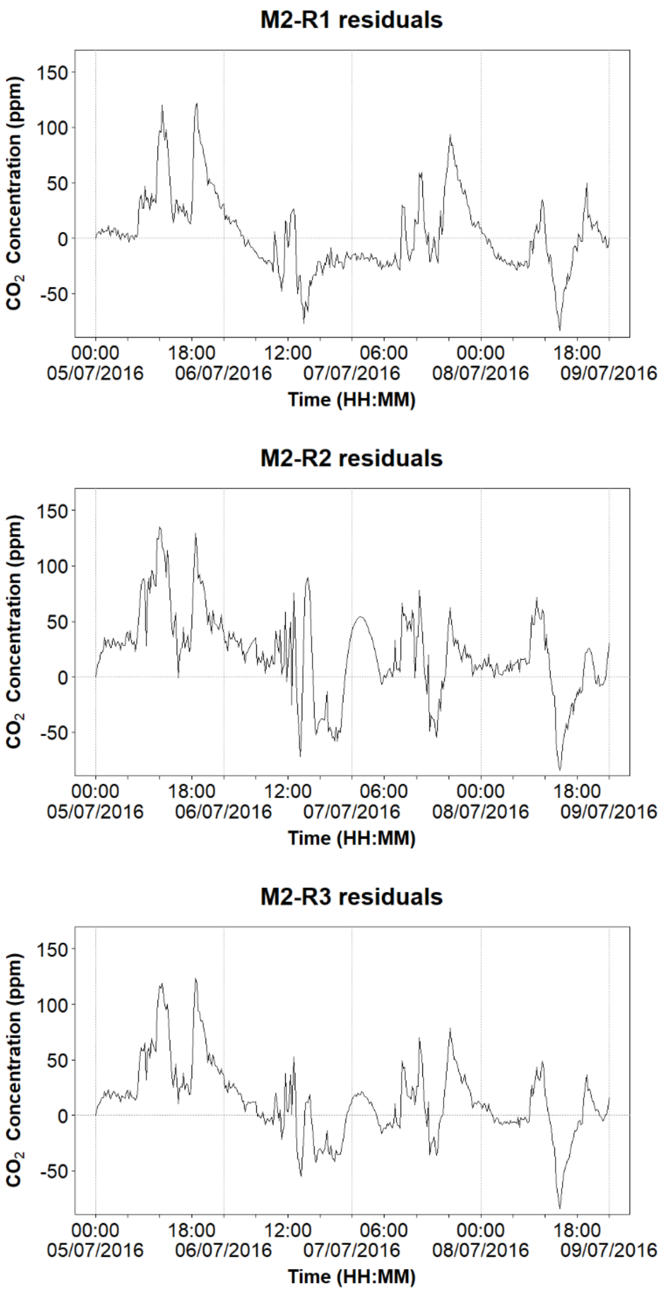

Fig. 6. Pure simulation residual plots for M2-R1, M2-R2 and M2-R3.

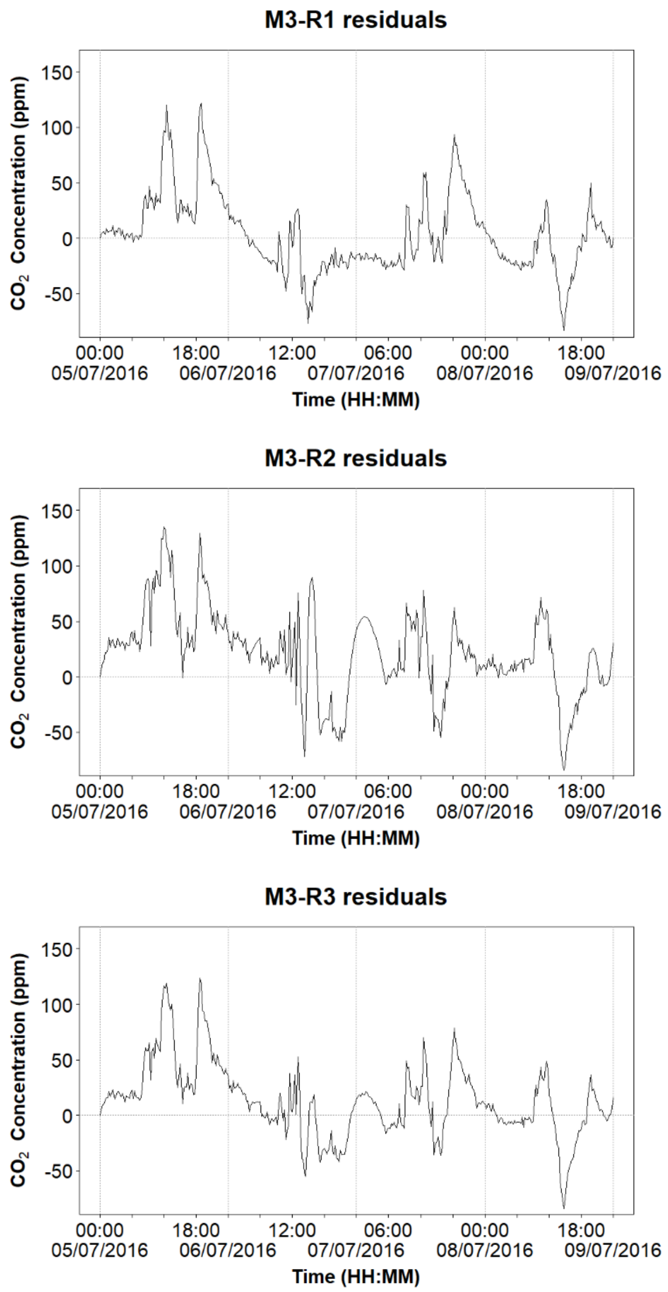

Fig. 7. Pure simulation residual plots for M3-R1, M3-R2 and M3-R3.

Although some lags fell outside the 95% confidence interval in all models and readings, they were quite small and were less than 5% (Fig. 8, Fig. 9 and Fig. 10). For this reason, it is reasonable to accept the hypothesis that there was no lag dependency in the one-step-ahead prediction residuals. The cumulated periodogram of M1-R1 fell slightly outside the 95% confidence interval (Fig. 8). However, M1-R2 and M1-R3 lay inside the confidence bands.

The cumulated periodograms of M2 (Fig. 9) were slightly better than the periodograms of M1. Only M2-R1 fell slightly outside the 95% confidence interval. Finally, there was no difference between the periodograms of M2 and M3 (Fig. 10). According to these results, we can affirm that all models were detailed enough to describe the CO_2 dynamics, and the one-step-ahead residuals obtained could be considered white noise. Consequently, the assumption of white noise residuals was fulfilled.

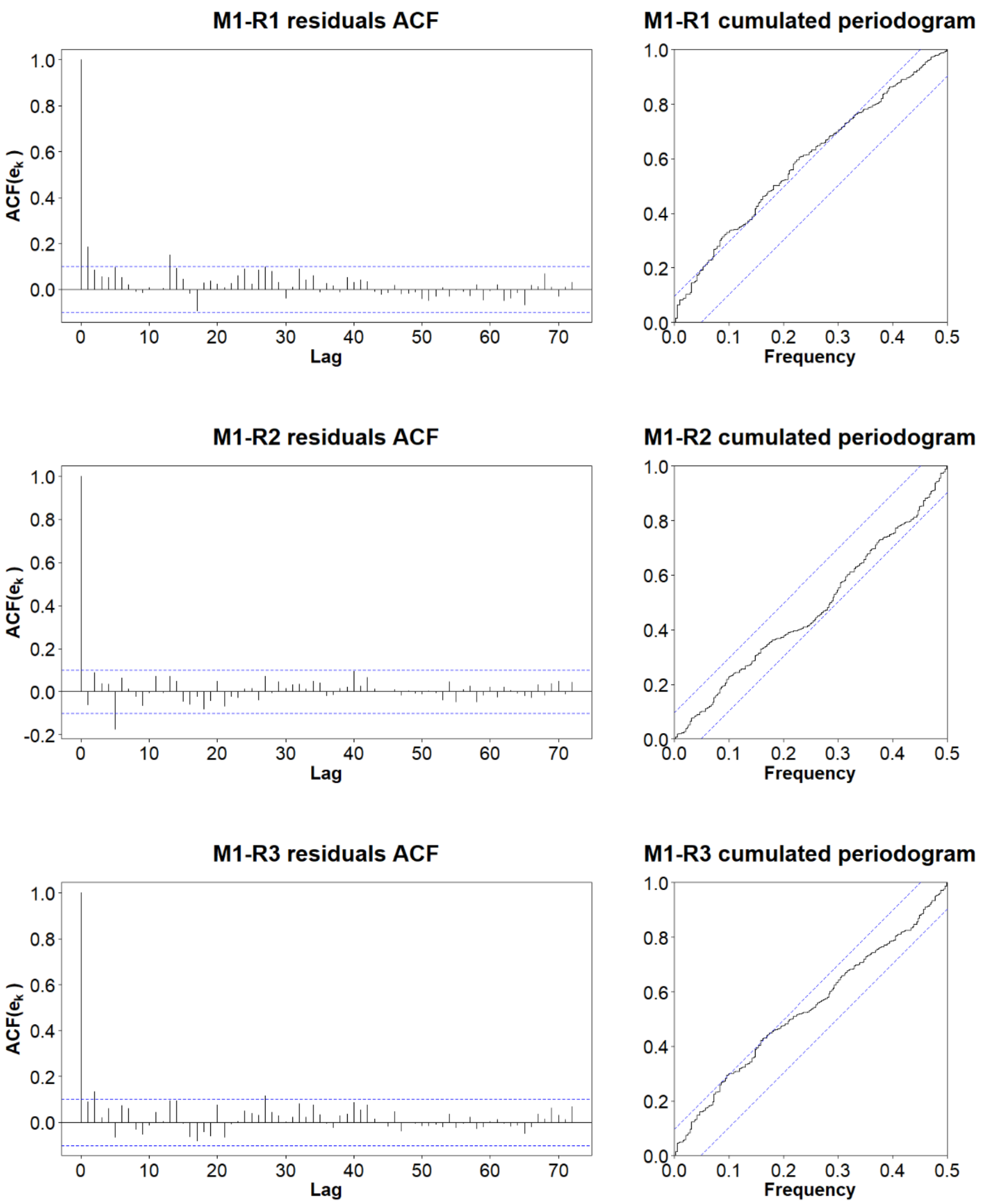

Fig. 8. The graphs on the left plot the autocorrelation function (ACF) for the residuals of M1-R1, M1-R2 and M1-R3; and the graphs on the right present the cumulated periodogram for M1-R1, M1-R2 and M1-R3

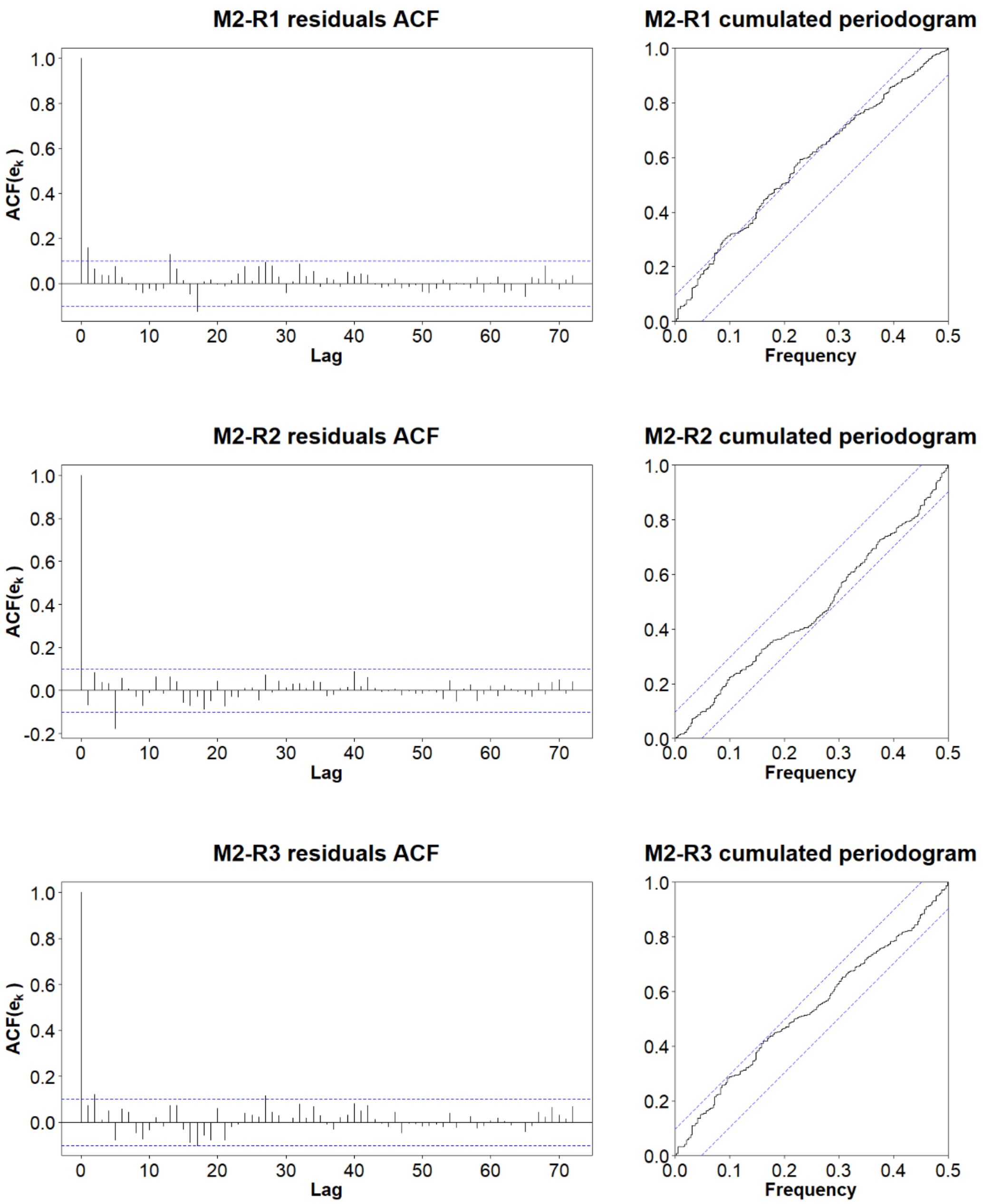

Fig. 9. The graphs on the left plot the autocorrelation function (ACF) for the residuals of M2-R1, M2-R2 and M2-R3; and the graphs on the right present the cumulated periodogram for M2-R1, M2-R2 and M2-R3.

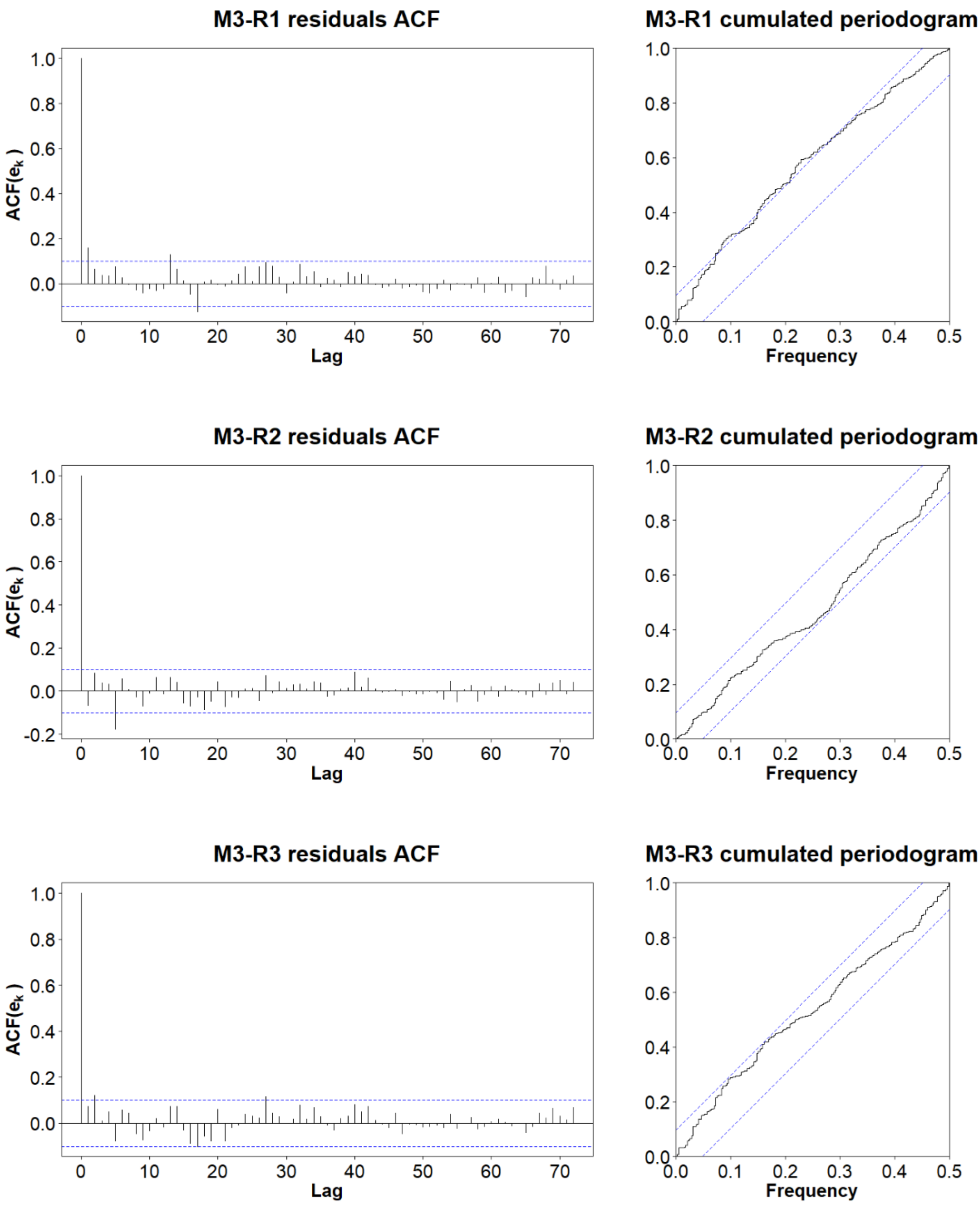

Fig. 10. The graphs on the left plot the autocorrelation function (ACF) for the residuals of M3-R1, M3-R2 and M3-R3; and the graphs on the right present the cumulated periodogram for M3-R1, M3-R2 and M3-R3.

The values of RMSE ranged from 35 to 45. The use of one sensor in the ventilation air flow reduced the RMSE (from 39-45 to 35-44, see Table 4). However, the use of one sensor outside the building to consider the window infiltrations did not improve the RMSE. The accuracy of all models was close to the accuracy of other studies that used stochastic methods to characterize the parameters of a ventilation system [19]. The RMSE values were slightly lower than in other studies that used deterministic approaches, and the reported RMSE ranged from 50 to 60 ppm [9]. The accuracy reported by M1, M2 and M3 was close to the accuracy of most commercial CO_2 sensors. In this research, the sensors' accuracy was ±2% of full scale. Consequently, the accuracy of the sensor was around ±60 ppm. Therefore, the accuracy reported by the three models and each position could be considered acceptable.

According to the results of all statistical tests presented for M1 and M2, we confirm that the models can represent the indoor CO_2 concentration over time. The use of one sensor in the ventilation air flow improved the performance of the models. However, when the variability of CO_2 concentration in the ventilation air flow was low, this sensor was not required. M3 should be discarded because it was over-parametrized.

All estimated parameters reported by M1, M2 and M3 were feasible in terms of the physics of the system. The estimations of the ventilation air flows (Q_{ven}) ranged between 28.77 m^3/h (0.30 h^{-1}) and 35.31 m^3/h (0.37 h^{-1}). These values are similar to the results of similar research in the field. Usually, research on natural ventilation in buildings with similar experimental conditions reports ratios of air change between 0.21 and 0.5 h^{-1} [15,19,24]. The ventilation air change rates that were obtained are lower than in other studies on forced ventilation, such as Amai and Novoselac [25], who reported air change rates between 1.1 and 8.7 h^{-1}.

The human emission rates of CO_2 (K_{occ}) reported in this research ranged between 13.56 L/h per person and 14.18 L/h per person. These values are close to previous studies that used stochastic methods to determine the parameters of a ventilation system (12.80 L/h) [19], and studies that used environmental chamber experiments to determine the CO_2 generation rate per person (12.60 L/h) [26]. However, the reported values were 21-24% lower than values assumed by other studies in the field of building ventilation with the same type of users (18.00-18.70 L/h) [16,27,28]. Generally, values used in the literature are calculated using the empirical equation for metabolic rate, the DuBois formula to calculate the nude body surface area, and the Nishi empirical formula. These formulas consider the size of the body and the level of physical activity.

The variability of Q_{ven} estimations was very low and ranged between 2.51% and 3.94% (see Table 4). The estimations of Q_{ven} were statistically different, because not all the 95% confidence intervals overlapped. However, the estimations of Q_{ven} among all models for the same position were not statistically different. In comparison, the variability of K_{occ} ranged between 1.84% and 3.07%. These values were slightly lower than the Q_{ven} variabilities. In addition, the estimation of K_{occ} was not statistically different, because all the 95% confidence intervals overlapped (Fig. 11).

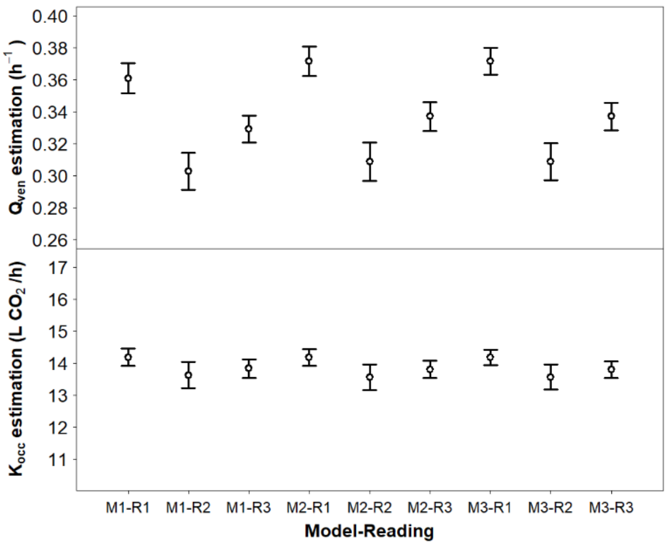

Fig. 11. Graphical representation of 95% confidence intervals for Q_{ven} and K_{occ} estimations

The mean of all Q_{ven} estimations was 31.96 m^3/h (0.34 h^{-1}), and the coefficient of variation was 8% (Table 5). The mean of all K_{occ} estimations was 13.86 L/h per person. Considering these results, it can be affirmed that the tested models are robust. This affirmation can be made because the estimation of K_{occ} was not statistically different between each model and position. Although the estimation of Q_{ven} was statistically different between each other, the variability in the estimations was low. This was an expected result because slightly different air movements could exist in the representative regions where a CO2 sensor should be installed [3].

Table 8. Estimation variability

Parameter	Q_{ven}	K_{occ}
Estimation mean	31.96 m^3/h (0.34 h^{-1})	13.86 L CO_2/h
Estimation standard deviation	2.56 m^3/h (0.03 h^{-1})	0.26 L CO_2/h
Estimation coefficient of variation	8.0 %	1.9%

4 Conclusion

This study presents how grey-box modelling can be used to estimate the natural ventilation in a room. In addition, it investigates the influence of sensor height when the above method is used. Two locations were selected to avoid the direct influence of room users. Three models based on stochastic differential equations were used to estimate the room ventilation air change rate using three different inputs: data from a sensor located on the ceiling, data from a sensor located at desk level, and the average data from both sensors. Then, the model parameters were identified using the maximum likelihood method. The models were validated using a set of statistical methods and physical interpretation of the estimated parameters.

Although the estimations of the ventilation air change rates were statistically different, the variability of the estimations was very low (coefficient of variation 8.0%). However, the estimations of the human emission rates of CO_2 were not statistically different among all models and readings. According to the results, when stochastic methods are used, the height of the sensor is less important as long as the direct user's influence is avoided. Maintenance teams tend to install building sensors high up where users cannot reach them, especially in buildings that have many users, such as schools or commercial centres. The aim of this practice is to avoid vandalism. Consequently, the proposed approach is compatible with current practice.

The tests were carried out in a room with natural ventilation. However, in rooms with forced ventilation, the results should be similar or even better. Forced ventilation helps to homogenise the air in a room. In other words, forced ventilation helps to achieve a perfectly mixed condition in a room. In this case, the location of CO_2 sensors is less important.

The results of this paper can be used for rooms of a similar height (up to 3 m). However further research must be carried out to test greater heights.

Another relevant result is that with the current CO_2 sensor technology available in the building market, the grey-box modelling approach cannot distinguish between natural ventilation flow from the corridor and natural ventilation flow from window infiltrations without additional experiments. Although the accuracy of most commercial CO_2 sensors is usually up to 60 ppm, the proposed approach provides a fast, cheap procedure for estimating room ventilation air change rate. Further research is required with more accurate sensors to determine whether this approach enables us to distinguish between natural ventilation air flow sources.

References

[1] A. Persily, Challenges in developing ventilation and indoor air quality standards: The story of ASHRAE Standard 62, Build. Environ. 91 (2015) 61–69. doi:10.1016/j.buildenv.2015.02.026.

[2] I. Škrjanc, B. Šubic, Control of indoor CO2 concentration based on a process model, Autom. Constr. 42 (2014) 122–126. doi:10.1016/j.autcon.2014.02.012.

[3] A. Bulińska, Z. Popiołek, Z. Buliński, Experimentally validated CFD analysis on sampling region determination of average indoor carbon dioxide concentration in occupied space, Build. Environ. 72 (2014) 319–331. doi:10.1016/j.buildenv.2013.11.001.

[4] S. Cui, M. Cohen, P. Stabat, D. Marchio, CO2 tracer gas concentration decay method for measuring air change rate, Build. Environ. 84 (2015) 162–169. doi:10.1016/j.buildenv.2014.11.007.

[5] N. Mahyuddin, H. Awbi, The spatial distribution of carbon dioxide in an environmental test chamber, Build. Environ. 45 (2010) 1993–2001. doi:10.1016/j.buildenv.2010.02.001.

[6] C.C. Menassa, N. Taylor, J. Nelson, A framework for automated control and commissioning of hybrid ventilation systems in complex buildings, Autom. Constr. 30 (2013) 94–103. doi:10.1016/j.autcon.2012.11.022.

[7] A.K. Duun-Henriksen, S. Schmidt, R.M. Røge, J.B. Møller, K. Nørgaard, J.B. Jørgensen, et al., Model Identification Using Stochastic Differential Equation Grey-

Box Models in Diabetes, J. Diabetes Sci. Technol. 7 (2013) 431–440. doi:10.1177/193229681300700220.

[8] M. Arnesano, G.M. Revel, F. Seri, A tool for the optimal sensor placement to optimize temperature monitoring in large sports spaces, Autom. Constr. 68 (2016) 223–234. doi:10.1016/j.autcon.2016.05.012.

[9] A. Pantazaras, S.E. Lee, M. Santamouris, J. Yang, Predicting the CO2 levels in buildings using deterministic and identified models, Energy Build. 127 (2016) 774–785. doi:10.1016/j.enbuild.2016.06.029.

[10] N. Mahyuddin, H.B. Awbi, M. Alshitawi, The spatial distribution of carbon dioxide in rooms with particular application to classrooms, Indoor Built Environ. 23 (2014) 433–448. doi:10.1177/1420326X13512142.

[11] A. Thavlov, H. Madsen, A non-linear stochastic model for an office building with air infiltration, Int. J. Sustain. Energy Plan. Manag. 7 (2015) 55–66. doi:10.5278/IJSEPM.2015.7.5.

[12] K.K. Andersen, H. Madsen, L.H. Hansen, Modelling the heat dynamics of a building using stochastic differential equations, Energy Build. 31 (2000) 13–24. doi:10.1016/S0378-7788(98)00069-3.

[13] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy Build. 43 (2011) 1511–1522. doi:10.1016/j.enbuild.2011.02.005.

[14] N.R. Kristensen, H. Madsen, S.B. Jørgensen, A method for systematic improvement of stochastic grey-box models, Comput. Chem. Eng. 28 (2004) 1431–1449.

doi:10.1016/j.compchemeng.2003.10.003.

[15] H. Li, X. Li, M. Qi, Field testing of natural ventilation in college student dormitories (Beijing, China), Build. Environ. 78 (2014) 36–43. doi:10.1016/j.buildenv.2014.04.009.

[16] W. Zhang, L. Wang, Z. Ji, L. Ma, Y. Hui, Test on Ventilation Rates of Dormitories and Offices in University by the CO2 Tracer Gas Method, in: Procedia Eng., Elsevier, 2015: pp. 662–666. doi:10.1016/j.proeng.2015.08.1061.

[17] F.D. Heidt, H. Werner, Microcomputer-aided measurement of air change rates, Energy Build. 9 (1986) 313–320. doi:10.1016/0378-7788(86)90036-8.

[18] D. Laussmann, D. Helm, Air Change Measurements Using Tracer Gases: Methods and Results. Significance of air change for indoor air quality, in: Chem. Emiss. Control. Radioact. Pollut. Indoor Air Qual., InTech, 2011. doi:10.5772/18600.

[19] M. Macarulla, M. Casals, M. Carnevali, N. Forcada, M. Gangolells, Modelling indoor air carbon dioxide concentration using grey-box models, Build. Environ. 117 (2017) 146–153. doi:10.1016/j.buildenv.2017.02.022.

[20] Scripps O2 program, CO2 Data, (2016). http://scrippso2.ucsd.edu/cosub2sub-data.

[21] N.R. Kristensen, H. Madsen, S.B. Jørgensen, Parameter estimation in stochastic grey-box models, Automatica. 40 (2004) 225–237. doi:10.1016/j.automatica.2003.10.001.

[22] CTSM-R Development Team, Continuous Time Stochastic Modeling in R User's Guide and Reference Manual, 2015. http://ctsm.info/.

[23] CTSM-R Development Team, Grey-box modeling of the heat dynamics of a building with CTSM-R, 2013. http://ctsm.info/.

[24] L.C. Ng, J. Wen, Estimating building airflow using CO2 measurements from a distributed sensor network, HVAC&R Res. 17 (2011) 344–365. doi:10.1080/10789669.2011.572223.

[25] H. Amai, A. Novoselac, Experimental study on air change effectiveness in mixing ventilation, Build. Environ. 109 (2016) 101–111. doi:10.1016/j.buildenv.2016.09.015.

[26] M.S. Zuraimi, A. Pantazaras, K.A. Chaturvedi, J.J. Yang, K.W. Tham, S.E. Lee, Predicting occupancy counts using physical and statistical Co 2 -based modeling methodologies, Build. Environ. 123 (2017) 517–528. doi:10.1016/j.buildenv.2017.07.027.

[27] D.A. Krawczyk, A. Rodero, K. Gładyszewska-Fiedoruk, A. Gajewski, CO 2 concentration in naturally ventilated classrooms located in different climates—Measurements and simulations, Energy Build. 129 (2016) 491–498. doi:10.1016/j.enbuild.2016.08.003.

[28] D.S. Dougan, L. Damiano, CO2-Based demand control ventilation. Do risks outweigh potential rewards?, ASHRAE J. (2004) 53.